# Real-Time Nowcasting the US Output Gap: Singular Spectrum Analysis at Work

**de Carvalho, M., and Rua, A. (2017), "Real-Time Nowcasting the US Output Gap: Singular Spectrum Analysis at Work," International Journal of Forecasting, 33, 185--198.**

# Real-time nowcasting the US output gap: Singular spectrum analysis at work

Miguel de Carvalho [a] and António Rua [b, c, *]

[a] *Faculty of Mathematics, Pontificia Universidad Católica de Chile, Santiago, Chile*

e-mail: MdeCarvalho@mat.puc.cl

[b] *Economics and Research Department, Banco de Portugal, Lisboa, Portugal*

[c] *Nova School of Business and Economics, Universidade Nova de Lisboa, Lisboa, Portugal*

e-mail: Antonio.Rua@bportugal.pt

**Abstract**

We explore a new approach for nowcasting the output gap based on singular spectrum analysis. Resorting to real-time vintages, a recursive exercise is conducted so to assess the real-time reliability of our approach for nowcasting the US output gap, in comparison with some well-known benchmark models. For our applied setting of interest, the preferred version of our approach consists of a multivariate singular spectrum analysis, where we use a Fisher $g$ test to infer which components, within the standard business cycle range, should be included in the grouping step. We find that singular spectrum analysis provides a reliable assessment of the cyclical position of the economy in real-time, with the multivariate approach outperforming substantially the univariate counterpart.

*Keywords*: Band-pass filter; Fisher $g$ test; Multivariate singular spectrum analysis; Singular spectrum analysis; US output gap, Real-time data.

*JEL classification*: C50, E32.

*Corresponding address: Economics and Research Department, Banco de Portugal, Av. Almirante Reis no. 71, 1150–012 Lisboa, Portugal.

# 1 Introduction

The output gap plays a central role in policymaking. Most central banks aim to keep inflation under control, and the output gap is a key source of inflation pressures in the economy. Given that the output gap fluctuates when the economy is overheating or underperforming, the conduct of monetary policy should take it into full consideration. It can also be used to determine and pursue policy measures by governments—as the cyclical position of the economy may influence fiscal policy—, and thus the assessment of the output gap is crucial for many formulations of countercyclical stabilization policy.

Measuring the output gap is however challenging—as it cannot be observed directly—and so cannot be assessed precisely. The revisions to which real-time output gap estimates are subject, present yet another challenge as they can compromise their operational usefulness for policymakers—who need reliable 'intel' in real-time. There are by now several studies documenting the large uncertainty of real-time output gap estimates, with this being a common issue for all estimation methods available (see Orphanides & van Norden, 2002, Orphanides, 2003a, Watson, 2007, Marcellino & Musso, 2011, Edge & Rudd, 2012, among others). The policy implications of the effects of output gap uncertainty have been addressed by, for example, Orphanides (2001), Rudebusch (2001), Smets (2002), Orphanides (2003b), and Orphanides and Williams (2007).

In this paper we focus on singular spectrum analysis (SSA), and evaluate its potential contribution for nowcasting output gap in a real-time setup. Despite the potential usefulness of SSA for the analysis of economic phenomena there are only a few applications in the economics and finance literature. In this respect, see the recent work by Hassani, Heravi, and Zhigljavsky (2009), Patterson, Hassani, Heravi, and Zhigljavsky (2011), de Carvalho, Rodrigues, and Rua (2012), Hassani, Heravi, and Zhigljavsky (2013), and Hassani, Soofi, and Zhigljavsky (2013).

The implementation of SSA involves the choice of two important parameters. In the decomposition stage one has to set the window length, $L$, in the embedding step, whereas in the reconstruction stage one has to choose the number of components for the grouping step. Concerning the choice of $L$, as pointed out by Hassani, Mahmoudvand, and Zokaei (2011), large values of $L$ allow longer period oscillations to be resolved, but choosing $L$ too large leaves too few observations from which to estimate the covariance matrix of the $L$ variables. It has been recommended that $L$ should be large

enough but not larger than $T/2$ (Golyandina, Nekrutkin, and Zhigljavsky, 2001), whereas Elsner and Tsonis (1996) discuss the practice of choosing $L$ equal to $T/4$, where $T$ is the sample size. Hassani, Heravi, and Zhigljavsky (2009) argue that if the time series presents a periodic component with an integer period, it is advisable to take the window length proportional to that period so to achieve a better separability of the periodic component. Drawing on the concept of separability between the signal and noise component, Hassani, Mahmoudvand, and Zokaei (2011) find that a suitable value for $L$, at least for reconstruction purposes, is close to Median$\{1, \ldots, T\}$ for a series of length $T$ (see also Hassani et al., 2012). However, as mentioned by Hassani, Heravi, and Zhigljavsky (2013), such value may not be optimal for forecasting purposes (Mahmoudvand, Najari, and Zokaei, 2013). In this respect, Hassani, Soofi, and Zhigljavsky (2013) take into account the forecasting horizon of interest when selecting $L$ whereas Hassani, Soofi, and Zhigljavsky (2015) choose the value of $L$ that minimizes the forecasting error. In the context of output gap estimation, de Carvalho, Rodrigues, and Rua (2012) suggest selecting $L$ as the maximum period of the business cycle frequency range one is interested in. As stressed by Hassani (2007), the selection of the proper window length depends upon the problem at hand, and on preliminary information about the time series. Finally, it should be noted that rules which may be optimal for (univariate) SSA need not to be optimal for multivariate SSA (MSSA). For example the above-mentioned rule based on Median$\{1, \ldots, T\}$ is suboptimal for MSSA; cf Hassani and Mahmoudvand (2013, Proposition 1).

The selection of the components to be used in the reconstruction stage is also far from straightforward. In the separation of signal and noise, one way to proceed is by looking at the plot of the eigenvalues, properly ordered by its value, associated with the reconstructed components. This plot can hopefully guide the truncation of the number of components to be considered in the grouping step. As mentioned by Hassani, Heravi, and Zhigljavsky (2009), a slowly decreasing sequence of eigenvalues is usually related to noise whereas similar values of the eigenvalues allow the identification of the eigentriples that correspond to the same harmonic component of the series. Furthermore, they suggest computing the periodogram for selecting the components to group. In fact, the presence of peaks in the periodogram provides an indication of the harmonic components in the series. Another possible approach is to compute $w$-correlations between the components; see Golyandina, Nekrutkin, and Zhigljavsky (2001) and Hassani (2007). Low values for $w$-correlations between reconstructed components

indicate that the components are well separated, whereas high values suggest that they should be considered as a group and possibly pertain to the same component in the SSA decomposition. Among other alternative approaches, one should mention Hassani, Soofi, and Zhigljavsky (2015) who choose the number of components that minimizes the forecasting error (see also Cassiano et al., 2013).

Recently, de Carvalho, Rodrigues, and Rua (2012) have shown that SSA can deliver output gap estimates that resemble those obtained with band-pass filters while improving the reliability of the corresponding nowcasts. Here we extend their work in several dimensions. First, to mimic a real-life policymaking scenario, that is, to replicate the problem faced by policymakers at the time policy decisions have to be taken, we consider real-time data. This means considering the vintages of data available at each moment in time. It is by now widely acknowledged that data revisions can affect policy decisions, and although the issue of the importance of data revisions is not recent, there has been a growing interest among practitioners to take on board real-time data into the analysis, since the influential work by Croushore and Stark (2001, 2003)—who compiled and examined real-time data for major US macroeconomic variables. Hence, we focus on the evaluation of output gap nowcasts computed through a recursive exercise using at each period the corresponding available vintage. This allows us to obtain real-time estimates—which are the ones relevant in terms of policymaking—whereas de Carvalho, Rodrigues, and Rua (2012) only considered *quasi*-real estimates, by considering the latest available vintage.[1]

Second, we suggest a novel approach for the selection of principal components to be used for reconstructing the cyclical component of a variable of interest. de Carvalho, Rodrigues, and Rua (2012) use an heuristic approach to select the components to be considered for the reconstruction of the cyclical component of GDP. Based on the dominant frequency, they consider the components that reflect periodicities of interest, namely within the business cycle frequency range. In this respect, Hassani, Heravi, and Zhigljavsky (2009) suggest the computation of the periodogram for assessing the dominant periodicity. We propose an alternative inferential procedure to address this issue, by using a spectral-based Fisher $g$ test. Although less popular than time domain analysis, Fourier analysis has

[1]The *quasi*-real estimate is the rolling estimate based on the latest available vintage. Both real-time and *quasi*-real estimates cover the same period, and they only differ due to data revisions. See Orphanides and van Norden (2002, p. 571).

proven to be quite useful in a wealth of contexts (see, for example, A'Hearn & Woitek, 2001, Rua & Nunes, 2005, Breitung & Candelon, 2006, Lemmens, Croux, & Dekimpe, 2008). Drawing on the periodogram estimator, Fisher (1929) derived an exact test—the so-called Fisher $g$ test—which allows for the detection of hidden periodicities of unspecified frequency, by determining whether a peak in the periodogram is significant or not. We use the Fisher $g$ test to select the principal components to be aggregated in the reconstruction of the output gap; specifically, we consider all principal components that present a statistically significant peak in the periodogram, within the standard business cycle frequency range. This provides a formal criterion for selecting the principal components relevant for the problem at hand. In this respect, we conduct a Monte Carlo simulation study to assess the behavior of the suggested procedure. Parenthetically, we note that Yarmohammadi (2011) also combine SSA with the Fisher $g$ test, but with a different goal in mind. Indeed, the author first suggests to preprocess the data using SSA and then applies the Fisher $g$ test with the aim of detecting hidden periodicities; thus for Yarmohammadi (2011), the goal is not on using the Fisher $g$ test to select the principal components to be aggregated in the reconstruction step, but rather to use SSA as a preprocessing step preliminarily to applying the Fisher $g$ test.

Another contribution of our paper rests on the use of information beyond that conveyed by GDP to estimate the output gap. Although, as stressed by Stock and Watson (1999), the cyclical component of real GDP is a useful proxy for the overall business cycle, it is sensible to argue that other macroeconomic variables should also reflect business cycle developments (see also the pioneer work of Burns and Mitchell, 1946). In this respect, the industrial production index is one of the macroeconomic indicators more commonly used in the literature for assessing the cyclical position of the economy in the absence of GDP data, and it is actually one of the top indicators used in practice for dating the US business cycle by the National Bureau of Economic Research (NBER) Business Cycle Dating Committee (`www.nber.org/cycles/recessions.html`). As GDP and industrial production are strongly correlated at business cycle frequencies, the use of industrial production data to complement GDP in the estimation of the output gap seems a natural choice. Typically, the use of macroeconomic data other than GDP does not lead to substantial differences in final output gap estimates, but can potentially improve the real-time assessment (see, for example, Valle e Azevedo, Koopman, & Rua, 2006, Valle e Azevedo, 2011). To take on board information beyond that conveyed by GDP, we extend the

output gap estimation from the SSA, considered in de Carvalho, Rodrigues, and Rua (2012), to the MSSA case.

To assess the relative performance of the suggested approach to nowcast the US output gap, we consider other econometric techniques, namely the popular Hodrick and Prescott (1997) filter and the band-pass filter of Christiano and Fitzgerald (2003).[2] In line with previous literature, we find that all approaches deliver relatively similar final output gap estimates. In addition, such estimates are in accordance with the US business cycle chronology. Based on a real-time US dataset and resorting to a standard battery of reliability statistics, we evaluate the real-time performance of each approach. The Hodrick–Prescott filter seems to perform the worst, whereas the SSA approach delivers more reliable output gap nowcasts than the alternative filtering techniques. Going beyond the univariate SSA, we conclude that the use of data other than GDP, in particular industrial production, can be very useful for improving output gap nowcasting. Hence, considering a multivariate framework based on SSA can be quite useful for producing reliable real-time estimates of the US output gap.

Our paper is organized as follows. In Section 2 we discuss our SSA-based approach for modeling business cycles. In Section 3 we conduct a Monte Carlo simulation study. In Section 4 we use our approach for real-time nowcasting the US output gap, and compare it with some popular benchmark methods. We conclude in Section 5.

# 2 Singular spectrum business cycle analysis

## 2.1 Modeling concept: Univariate setting

As argued by Morley and Piger (2012) there are two main views for modeling business cycles: An alternating-phases approach (Mitchell, 1927), which considers a rotating sequence of expansions–recessions, and an output gap approach (Beveridge & Nelson, 1981) where the business cycle, $c_t$, is defined as a transitory deviation from a trend, $\tau_t$. Formally, for seasonally adjusted data, the latter

[2]We underscore that although measuring output gap nowcast uncertainty can also be interesting (Garratt, Mitchell, and Vahey, 2014), the focus here is on point estimation.

approach is based on decomposing GDP, $y_t$, as follows

$$y_t = \tau_t + c_t. \tag{1}$$

The target of estimation in an output gap approach is thus naturally $C_t$. Since most business cycle literature has been concerned with recurring movements ranging from 6 to 32 quarters, a more reasonable working assumption is provided by the model

$$y_t = \tau_t + c_t + \varepsilon_t, \tag{2}$$

where $\varepsilon_t$ is a noise term describing recurring movements of frequencies higher than the ones of interest in a business cycle context. The SSA-based approaches to be discussed in the next sections are based on the output gap approach in Eq. (2), and the interest is on assessing the performance of the methods in real-time, so that our goal is on nowcasting $c_t$, using real-time data. In Fig. 1 we depict the vintage release schedule and corresponding period under analysis. The vintage at time $t$ includes the first estimate for time $t$; the vintage at time $t+1$ includes the first estimate for time $t+1$, and the second estimate for time $t$; the vintage at time $t+2$ includes the first estimate for time $t+2$, the second estimate for time $t+1$, and the third estimate for time $t$; and so on.

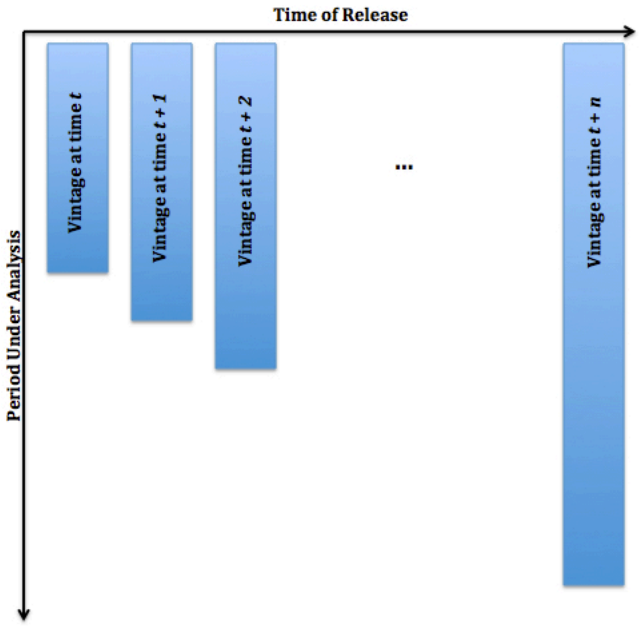



Figure 1: Vintage release schedule and corresponding periods under analysis.

For a primer on SSA see, for instance, Golyandina, Nekrutkin, and Zhigljavsky (2001), Hassani (2007), and Sanei and Hassani (2015). Below, we focus on discussing SSA in the context of our applied econometric problem of interest, so that the expression 'singular spectrum business cycle analysis' should be understood as a synonym of an adapted SSA with business cycle applications in mind.

## 2.2 Univariate singular spectrum business cycle analysis

The method entails two phases, namely decomposition and reconstruction, and each of these phases includes two steps; the phase of decomposition includes the steps of embedding and singular value decomposition, which we discuss below. Let $Y = (y_1, \ldots, y_T)$ denote a univariate time series from which we intend to extract information on the output gap (say GDP).

**Embedding.** This is the preliminary step of the method. SSA starts by organizing the original time series of interest, $Y$, into a so-called trajectory matrix, i.e., a matrix whose columns consist of rolling windows of length $L$, i.e.

$$\boldsymbol{X} = [X_1 \ \cdots \ X_K] = (x_{i.j})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_T \end{pmatrix}. \tag{3}$$

Here, $L$ denotes the window length, and $K = T - L + 1$. Since all elements over the diagonal $i+j = const$, are equal, the trajectory matrix $\boldsymbol{X}$ is a Hankel matrix.

**Singular Value Decomposition.** In the second step we perform a singular value decomposition of the trajectory matrix. Let $\lambda_1, \ldots, \lambda_L$ denote the eigenvalues of $\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}$, presented in decreasing order, and let $U_1, \ldots, U_L$ denote the corresponding eigenvectors. In this step, we decompose the trajectory matrix $\boldsymbol{X}$ as follows

$$\boldsymbol{X} = \sum_{i=1}^{d} \boldsymbol{X}_i, \tag{4}$$

where $\boldsymbol{X}_i = \sqrt{\lambda_i} U_i V_i^{\mathrm{T}}$, $V_i = \boldsymbol{X}^{\mathrm{T}} U_i / \sqrt{\lambda_i}$, and $d = \max\{i \in \{1, \ldots, L\} : \lambda_i > 0\}$.

Below we discuss the second phase of the method—reconstruction, which entails the steps of grouping cyclical components and diagonal averaging.

**Grouping Cyclical Components.** Not all summands in Eq. (4) contain relevant information on the business cycle,[3] and hence we confine ourselves to a subset $S_{\text{Fisher}}$ of $\{1, \ldots, d\}$, so to compute what we call the cycle matrix,

$$\boldsymbol{C} = (c_{i.j}^{S_{\text{Fisher}}})_{i,j=1}^{L,K} = \sum_{s \in S_{\text{Fisher}}} \boldsymbol{X}_s, \tag{5}$$

In practice, we construct $S_{\text{Fisher}}$ through a Fisher $g$ test on which we provide further details in Section 2.5.[4]

**Diagonal Averaging.** In this step we average over all the elements of the diagonal $i + j = const$ of the cycle matrix in Eq. (5) so to obtain a Hankel matrix, from where our business cycle indicator, $\widehat{C} = (\widehat{c}_1, \ldots, \widehat{c}_T)$, can be constructed. Essentially, our business cycle indicator is constructed by averaging the cycle matrix over the 'antidiagonals' $i + j = k + 1$. For example: $k = 1$, yields $\widehat{c}_1 = c_{1,1}^{S_{\text{Fisher}}}$; $k = 2$, yields $\widehat{c}_2 = (c_{1,2}^{S_{\text{Fisher}}} + c_{2,1}^{S_{\text{Fisher}}})/2$; $k = 3$, yields $\widehat{c}_3 = (c_{1,3}^{S_{\text{Fisher}}} + c_{2,2}^{S_{\text{Fisher}}} + c_{3,1}^{S_{\text{Fisher}}})/3$; *etc.*

## 2.3 Modeling concept: Multivariate setting

The method in Section 2.2 is essentially an updated version of the approach in de Carvalho, Rodrigues, and Rua (2012), but with one important difference: It includes a formal inference step based on Fisher $g$ test (Section 2.5) for selecting relevant business cycle components. We now generalize the approach to the multivariate setting. The main motivation for this is as follows: From a practical viewpoint, we have reasons to believe that we should be able to borrow strength from further information available

[3] Recall that most business cycle literature has been concerned with recurring movements ranging from 6 to 32 quarters.

[4] Often in SSA applications, the objective of the grouping step is on disentangling signal from noise, and in such case we typically write $\boldsymbol{X} = \sum_{i \in I} \boldsymbol{X}_i + \sum_{i \notin I} \boldsymbol{X}_i$, with the components in $I \subset \{1, \ldots, d\}$ representing the signal; criteria such as the ratio $\lambda_i / \sum_{j=1}^{d} \lambda_j$ are then often used to guide on the selection of $I$. In our context, criteria such as this one fail to provide a complete portrait, because we are only interested in the part of the signal associated with regular movements within certain (business cycle) frequencies, and this is one of the reasons why the Fisher $g$ test becomes an important tool in our setup.

on real time. Particularly, we are interested in constructing a business cycle indicator which combines information of the GDP and the industrial production (IP) index—which is a proxy for measuring economic activity evolution, and it is well known to be strongly correlated with the aggregate activity as measured by GDP (see, for instance, Fagiolo, Napoletano, & Roventini, 2008, de Carvalho & Rua, 2014). With this in mind, we extend the working assumption in (2) to a joint setting, so that the dynamics governing GDP, $y_t^{(1)}$, and other time series which we believe to be informative on the business cycle, $c_t^{(1)}$, such as IP, is the following

$$\begin{cases} y_t^{(1)} = \tau_t^{(1)} + c_t^{(1)} + \varepsilon_t^{(1)}, \\ \vdots \\ y_t^{(M)} = \tau_t^{(M)} + c_t^{(M)} + \varepsilon_t^{(M)}. \end{cases} \tag{6}$$

We underscore that the target of estimation is $C_t^{(1)}$, and thus the same as in Section 2.1.

## 2.4 Multivariate singular spectrum business cycle analysis

Multivariate singular spectrum business cycle analysis can be conducted by extending the approach discussed in Section 2.2. Suppose that we observe $M$ time series with possibly different lengths $Y^{(i)} = (y_1^{(i)}, \ldots, y_{T_i}^{(i)})$, for $i = 1, \ldots, M$. Multivariate singular spectrum business cycle analysis entails the following steps.

**Embedding.** Embedding transforms the time series $Y^{(i)}$ into a matrix $\boldsymbol{X}^{(i)} := [X_1^{(i)} \ \cdots \ X_{K_i}^{(i)}]$, where $X_j^{(i)} = (y_j^{(i)}, \ldots, y_{j+L_i+1}^{(i)})^{\mathrm{T}}$, $L_i$ is the window length for each series, and $K_i = T_i - L_i + 1$. The matrix $\boldsymbol{X}^{(i)}$ is the so-called trajectory matrix, and it is a Hankel matrix. The outcome of this step is a block Hankel matrix, $\boldsymbol{X}_V$, which can be constructed if $K_1 = \cdots = K_M = K$ (but with possibly different $L_i$ and $T_i$), by setting

$$\boldsymbol{X}_V = \begin{bmatrix} \boldsymbol{X}^{(1)} \\ \vdots \\ \boldsymbol{X}^{(M)} \end{bmatrix}. \tag{7}$$

The subscript '$V$' in $\boldsymbol{X}_V$ is added to highlight that this is a block Hankel trajectory matrix constructed in the vertical form. Another alternative block Hankel matrix, $\boldsymbol{X}_H$, could be constructed if $L_1 = \cdots =$

$L_M = L$ (but with a possibly different $K_i$ and $T_i$), by considering

$$\boldsymbol{X}_H = [\boldsymbol{X}^{(1)} \ \cdots \ \boldsymbol{X}^{(M)}]. \tag{8}$$

Motivated by theoretical and empirical considerations in Hassani and Mahmoudvand (2013, Sections 4–5), throughout we focus on the vertical form (7). Particularly, we underscore that

$$\mathbf{X}_H \mathbf{X}_H^{\mathrm{T}} = \boldsymbol{X}^{(1)} \boldsymbol{X}^{(1)^{\mathrm{T}}} + \cdots + \boldsymbol{X}^{(M)} \boldsymbol{X}^{(M)^{\mathrm{T}}},$$

and thus the horizontal form does not take into account the cross terms $\boldsymbol{X}^{(i)} \boldsymbol{X}^{(j)^{\mathrm{T}}}$(which can be related to the different pairs of time series $Y^{(i)}$ and $Y^{(j)}$, for $i \neq j$), whereas the vertical form incorporates them into the analysis (cf. Eq. (10) below).

**Singular Value Decomposition.** In the second step we perform a singular value decomposition of $\boldsymbol{X}_V$. Let $\lambda_1, \ldots, \lambda_{L_{\mathrm{sum}}}$ denote the eigenvalues of $\boldsymbol{X}_V \boldsymbol{X}_V^{\mathrm{T}}$, presented in decreasing order, and let $U_1, \ldots, U_{L_{\mathrm{sum}}}$ denote the corresponding eigenvectors, with $L_{\mathrm{sum}} = \sum_{i=1}^{M} L_i$. Thus, we decompose the trajectory matrix $\boldsymbol{X}_V$ into

$$\boldsymbol{X}_V = \sum_{i=1}^{L_{\mathrm{sum}}} \boldsymbol{X}_i, \tag{9}$$

where $\boldsymbol{X}_i = \sqrt{\lambda_i} U_i V_i^{\mathrm{T}}$, and $V_i = \boldsymbol{X}_V^{\mathrm{T}} U_i / \sqrt{\lambda_i}$. Note that

$$\boldsymbol{X}_V \boldsymbol{X}_V^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{X}^{(1)} \boldsymbol{X}^{(1)^{\mathrm{T}}} & \boldsymbol{X}^{(1)} \boldsymbol{X}^{(2)^{\mathrm{T}}} & \cdots & \boldsymbol{X}^{(1)} \boldsymbol{X}^{(M)^{\mathrm{T}}} \\ \boldsymbol{X}^{(2)} \boldsymbol{X}^{(1)\mathrm{T}} & \boldsymbol{X}^{(2)} \boldsymbol{X}^{(2)^{\mathrm{T}}} & \cdots & \boldsymbol{X}^{(2)} \boldsymbol{X}^{(M)^{\mathrm{T}}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{X}^{(M)} \boldsymbol{X}^{(1)^{\mathrm{T}}} & \boldsymbol{X}^{(M)} \boldsymbol{X}^{(2)^{\mathrm{T}}} & \cdots & \boldsymbol{X}^{(M)} \boldsymbol{X}^{(M)^{\mathrm{T}}} \end{bmatrix}, \tag{10}$$

and hence the matrices $\boldsymbol{X}^{(i)} \boldsymbol{X}^{(i)^{\mathrm{T}}}$ corresponding to SSA appear in the diagonal of the block matrix $\boldsymbol{X}_V \boldsymbol{X}_V^{\mathrm{T}}$.

**Grouping Cyclical Components.** Not all summands in Eq. (9) contain relevant information on the business cycle, and hence we confine ourselves to a subset $S_{\mathrm{Fisher}}$ of $\{1, \ldots, L_{\mathrm{sum}}\}$, so to produce

what we call the cycle matrix

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}^{(1)} \\ \vdots \\ \boldsymbol{C}^{(M)} \end{bmatrix} = \begin{bmatrix} (c_{i.j}^{(1)})_{i,j=1}^{L_1,K} \\ \vdots \\ (c_{i.j}^{(M)})_{i,j=1}^{L_M,K} \end{bmatrix} = \sum_{s \in S} \boldsymbol{X}_s. \tag{11}$$

Similarly to Section. 2.2, we construct $S$ through a Fisher $g$ test on which we provide further details in Section 2.5.[5] As we discuss in Section 4 the advantages of our Fisher–$g$ test approach are particularly evident in the multivariate setting, given that we face a larger number of 'candidate' components which could potentially be used to construct the cycle.

**Diagonal Averaging.** Our business cycle indicator, $\widetilde{C} = (\tilde{c}_1, \ldots, \tilde{c}_T)$, is finally given by the diagonal averaging of $\boldsymbol{C}^{(1)}$, the block of the cycle matrix corresponding to GDP, along the same lines as discussed above. Throughout we follow the convention of using tildes to denote the business cycle indicator yield through MSSA ($\widetilde{C}$) and hats for SSA ($\widehat{C}$).

## 2.5 Targeted grouping based on the Fisher $g$-statistic

The grouping stage in SSA should take into account the targeted output. In our framework, the aim is to group the components that reflect business cycle developments. In this respect, de Carvalho, Rodrigues, and Rua (2012) have grouped the components that seemed, by visual inspection, to contain information about the standard business cycle frequency range. Here, we suggest a formal inferential approach to address this issue. Underlying the informal approach of de Carvalho, Rodrigues, and Rua (2012) is the idea that one should select the components whose dominant periodicity (or frequency) falls within the range of frequencies of interest. This problem can be more formally addressed using spectral analysis. In particular, one can determine the dominant frequency (or periodicity) by finding the peak in the periodogram, while its statistical significance can be assessed through the so-called Fisher $g$-statistic—to be introduced below. If the frequency at which the peak is observed in the periodogram lies within the business cycle frequency range, and if it is statistically significant according to Fisher $g$-statistic, then that component is selected for the reconstruction of the cyclical component.

[5] For ease of notation we drop the dependence of $c_{i.j}$ on $S_{\text{Fisher}}$ in (11) and throughout the remainder of the paper.

As mentioned earlier, the Fisher $g$ test draws on the periodogram; see Priestley (1981, Sec. 6.1.4). The periodogram unveils the power of the signal at various frequencies, so that if the signal is being driven by a certain frequency, the periodogram presents a peak at that periodicity. Basically, the Fisher $g$ test checks for the proportion of power accounted for the frequency associated with the peak in the periodogram, and tests whether such peak is random or not. More formally, if $Y = (y_1, \ldots, y_T)$ is an equally-spaced time series, the periodogram consists of the set of points

$$\{(\omega_j, I(\omega_j)) : j = 1, \ldots, J\}, \quad J = \lfloor (T-1)/2 \rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the floor function, $\omega_j = 2\pi j/T$ are the so-called Fourier frequencies, for $j = 1, \ldots, J$, and

$$I(\omega) = \frac{1}{T}\left|\sum_{t=1}^{T} y_t \mathrm{e}^{-\mathrm{i}\omega t}\right|^2 = \frac{1}{T}\left[\left\{\sum_{t=1}^{T} y_t \sin(\omega t)\right\}^2 + \left\{\sum_{t=1}^{T} y_t \cos(\omega t)\right\}^2\right], \quad \omega \in (0, \pi).$$

If a time series has a significant periodic component with frequency $\omega^\star$, then the periodogram will exhibit a peak at frequency $\omega^\star$. Fisher (1929), in a celebrated paper, derived an exact test for testing the significance of the spectral peak based on the $g$-statistic,

$$g = \frac{\max\{I(\omega_1), \ldots, I(\omega_J)\}}{\sum_{j=1}^{J} I(\omega_j)}, \tag{12}$$

In Fisher's test, the null hypothesis is that the spectral peak is not statistically significant against the alternative hypothesis that there is a periodic component; under the Gaussian assumption, large values of $g$ lead to the rejection of the null hypothesis. The $p$-value of the test under the null hypothesis is given by

$$p := P(g > g^\star) = \sum_{j=1}^{J} (-1)^{j-1} \frac{J!}{j!\,(J-j)!} (1 - jg^\star)_+^{J-1}, \tag{13}$$

where $x_+ = \max(x, 0)$ and $g^\star$ is the observed value of the $g$-statistic (cf. Brockwell and Davis, 1991, Corollary 10.2.2). In practice we proceed as in the following pseudocode implementation. Let $\Omega \subseteq (0, \pi)$ denote a range of frequencies of interest, and let $\overline{\mathbb{D}}_i = \overline{\mathbb{D}}(\sqrt{\lambda_i} U_i V_i^{\mathrm{T}})$ denote the $i$th principal component.

**Targeted grouping based on the Fisher $g$-statistic**

Start with $S^{(0)} = \emptyset$, and `for` $i = 1, \ldots, L_{\mathrm{sum}}$, `do`:

Step 1. Obtain the periodogram of $\overline{\mathbb{D}}_i$, here denoted as $I_i(\omega)$, and compute:

$$\omega_i^\star = \arg \max_{\omega \in \{\omega_1, \ldots, \omega_J\}} I_i(\omega). \tag{14}$$

Step 2. `If` $\omega_i^\star \in \Omega$ go to Step 3; otherwise increment $i$ and go back to Step 1.

Step 3. Use Eq. (12) to compute the $g$-statistic associated with $\overline{\mathbb{D}}_i$; save the result in $g_i$.

Step 4. Use Eq. (13) to compute the $p$-value corresponding to the $g_i$ statistic from Step 3; save the result in $p_i$.

Step 5. `If` $p_i < \alpha$, set $S^{(i)} = S^{(i-1)} \cup \{i\}$; otherwise, set $S^{(i)} = S^{(i-1)}$.

Step 6. `If` $i = L_{\text{sum}}$, set $S_g = S^{(i)}$ and `stop`; otherwise, increment $i$ and go back to Step 1.

Following the notation from the pseudocode implementation above, throughout we use the notations $g_i$ denote the Fisher $g$-statistic computed from the periodogram of $\overline{\mathbb{D}}_i$; similarly, $p_i$ is used to denote the $p$-value corresponding to this statistics, while $S_g$ denotes the grouping set selected through our approach. To be able to visualize in a simple way which components have been selected through our method, we propose plotting

$$\{(i, \delta_i(S_g)) : i = 1, \ldots, L_{\text{sum}}\}, \tag{15}$$

where $\delta.(\cdot)$ denotes the Dirac measure, and $S_g = \{i \in \{1, \ldots, L_{\text{sum}}\} : \omega_i^\star \in \Omega, p_i < \alpha\}$, where $\alpha$ denotes the significance level used to conduct the analysis; throughout we will call the graph in Eq. (15) as the *comb-plot*, and the point masses $\delta_i(S_g)$ as *Fisher g indicators*, for $i = 1, \ldots, L_{\text{sum}}$. Throughout we always report results for a significance level of 5%.

# 3 Simulation study

To assess the behavior and sensitivity of the proposed approach for extracting the cyclical component, we conduct a Monte Carlo simulation study. For this purpose we consider the case where a time series can be decomposed into a trend, a cyclical component, and an irregular term as in (2). In particular, we consider as data generating process the following time series model

$$y_t^{(1)} = t + \cos(\omega t) + \varepsilon_t,$$

where $0 < \omega < \pi$, $t = 1, \ldots, T$, and $\varepsilon_t$ is Gaussian white noise process with zero mean and variance $\sigma_\varepsilon^2$. The first term allows for a time trend, the second term generates a well-defined cyclical component with frequency $\omega$, and the last term is a noise sequence. Based on our dataset, we consider for the baseline specification $T = 250$ (which is close to our sample size), $L = 32$ (see de Carvalho, Rodrigues, and Rua, 2012), and we set $\sigma_\varepsilon^2$ such that the signal-to-noise ratio is 4, that is, the standard deviation of the cyclical component is four times the standard deviation of the irregular term, as found for the US GDP series.

The simulation exercise was conducted as follows. For each frequency, $\omega$, the time series is generated and then the above described filtering procedure is applied. The Mean Squared Error (MSE) between the estimated cyclical component and the true one is computed. Finally, the average MSE across $B = 1000$ simulations is calculated. Since we are interested in assessing performance over the spectrum of frequencies, we compute the average MSE for each frequency $\omega$, i.e.

$$\text{MSE}_\omega = \frac{1}{TB} \sum_{t=1}^{T} \sum_{b=1}^{B} \{\widehat{c}_{t,b} - \cos(\omega t)\}^2, \quad \omega \in (0, \pi), \tag{16}$$

where $\widehat{c}_{t,b}$ is the SSA business cycle indicator, corresponding to the $b$th simulated data set. The average MSE for each frequency $\omega$ is displayed in Fig. 2.

The baseline case is shown in Fig. 2(a) and corresponds to the bold line. For comparison, we also report the results of the well-known and commonly applied filters, proposed by Christiano and Fitzgerald (2003) and Hodrick and Prescott (1997). For these, we consider the usual parameter values to extract the GDP cyclical component. In particular, for the Hodrick–Prescott filter we set the smoothing parameter equal to 1600 which is the recommended value for quarterly data; see Prescott (1986) and also Baxter and King (1999) for a more thorough discussion. In the case of the Christiano–Fitzgerald filter we define the range of periodicities of interest, when extracting the GDP cyclical component, to be between 6 and 32 quarters which corresponds to the standard frequency range considered in the business cycle literature; see, for example, Stock and Watson (1999, 2005).

One can see from Fig. 2(a) that when the cycle is driven by a frequency within the frequency range of interest namely, between 6 and 32 time periods, the MSE is close to zero. As expected, when the frequency is outside this interval the SSA based approach does not retain any cyclical component and the MSE clearly departs from zero. The results obtained with SSA are very similar to the ones obtained

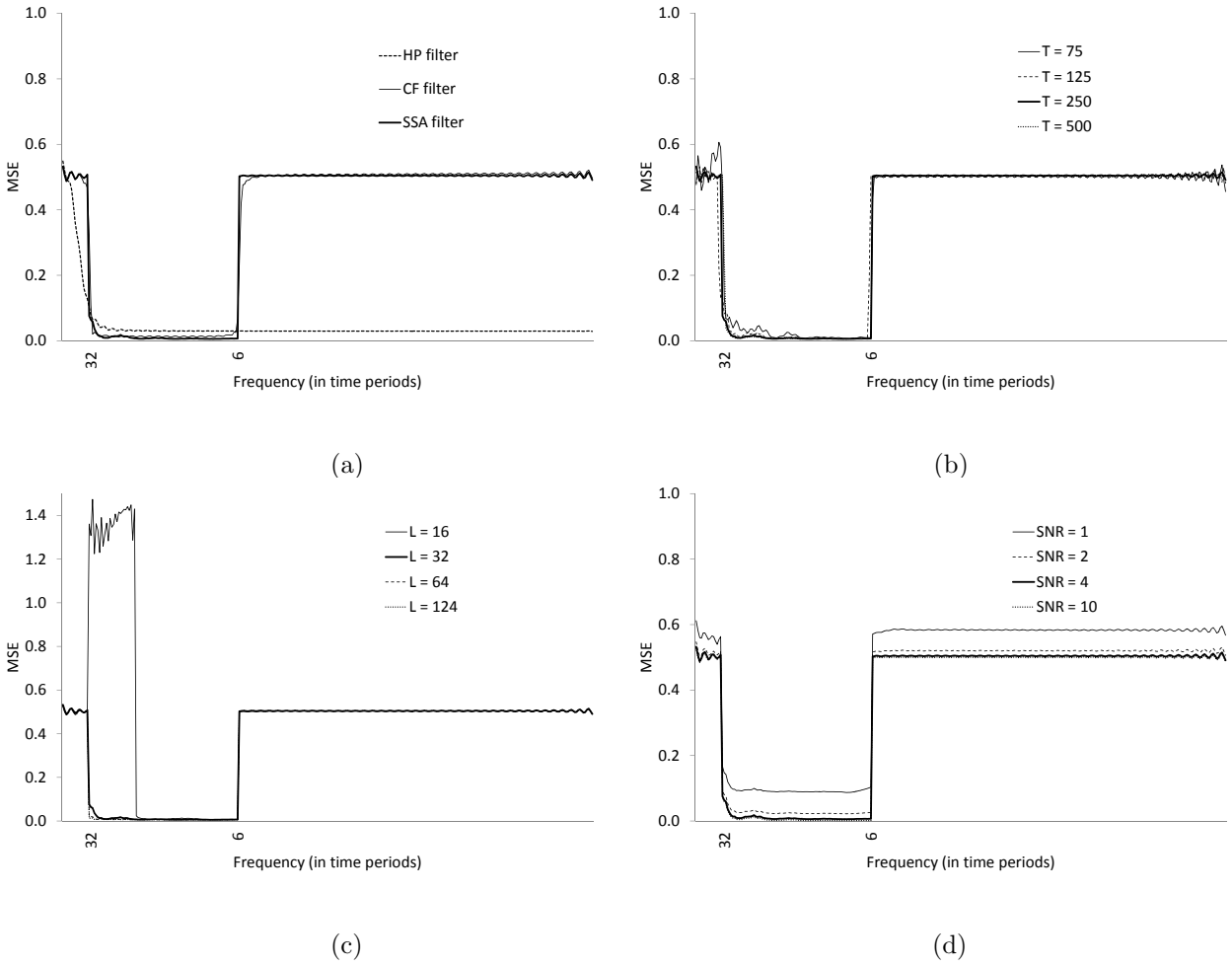



Figure 2: Simulation results on the univariate setting. MSE (Mean Square Error) over the spectrum, computed according to Eq. (16). Throughout all figures the bold line is used to represent the baseline specification with $T = 250$, $L = 32$, and $\sigma_\varepsilon^2$ such that the signal-to-noise ratio is 4. HP and CF respectively denote Hodrick–Prescott and Christiano–Fitzgerald, $T$ is the sample size, $L$ is the window length, and SNR stands for Signal-to-Noise Ratio.

with the Christiano–Fitzgerald filter which is intended to approximate an optimal band pass-filter. These findings reinforce the usefulness of the proposed method as a band-pass filtering procedure. In the case of the Hodrick–Prescott filter, higher frequencies are also retained in the estimated cyclical component (as seen from its MSE staying low for higher frequencies) reflecting the fact that the Hodrick–Prescott filter acts as a high-pass filter (King & Rebelo, 1993, Baxter & King, 1999).

We now assess the sensitivity of the results to different alternative specifications. Firstly, we consider several sample sizes, namely $T = 75$, 125, 250, 500. From Fig. 2(b) it can be observed that with moderate sample sizes the behavior is close to the one obtained with larger samples whereas it deteriorates a bit with small samples in particular for longer cycles. Take the case where $T = 75$. Naturally, it is hard to infer about cycles around 32 periods when one only observes 75 time periods.

Let us now consider different window lengths. Following de Carvalho, Rodrigues, and Rua (2012), $L = 32$ is used as baseline case. Alternatively, one can consider a value close to the Median$\{1, \ldots, T\}$ such as 124 (see Hassani, Mahmoudvand, and Zokaei, 2011). Hassani, Heravi, and Zhigljavsky (2009) argue that if the time series presents a periodic component with an integer period, it is advisable to take the window length proportional to that period so to achieve a better separability of the periodic component. Hence, and to take on board a value between 32 and 124 we consider $L = 64$. Note that both $L = 32$ and $L = 64$ are proportional to the maximum period of the frequency range of interest. In addition, we also consider a value lower than the baseline, $L = 16$. We find that the differences are negligible for $L = 32$, $64, 124$; see Fig. 2(c). However, when one considers a window length of size lower than the maximum period one is interested in there is a sharp deterioration of the MSE in the frequency range delimited by the maximum period of interest and $L$. For instance, if one sets $L = 16$, then the cyclical component is poorly estimated for cycles with a period between 16 and 32. Hence, based on this result, it is advisable to always choose an $L$ equal or larger than the maximum period of the business cycle frequency range of interest.

Finally, we also considered different signal-to-noise ratios (SNR) to assess the impact of changing the relative importance of the cyclical component vis-à-vis the noise. In the baseline specification this ratio has been set to 4 which is the value we find for the US GDP series. In addition, we consider a higher value, 10, as well as lower signal-to-noise ratios, 2 and 1. Compared to the baseline setting, the differences are negligible when one considers a noise reduction (see Fig. 2(d)). As expected, adding

more noise leads to a deterioration of the MSE. However, when the SNR is set to 2 it does not worsen too much the results. Only when we set the SNR to 1, there is a more visible deterioration of the MSE, but in this case the noise component is as important as the cyclical component which means in practice that the series is very noisy.

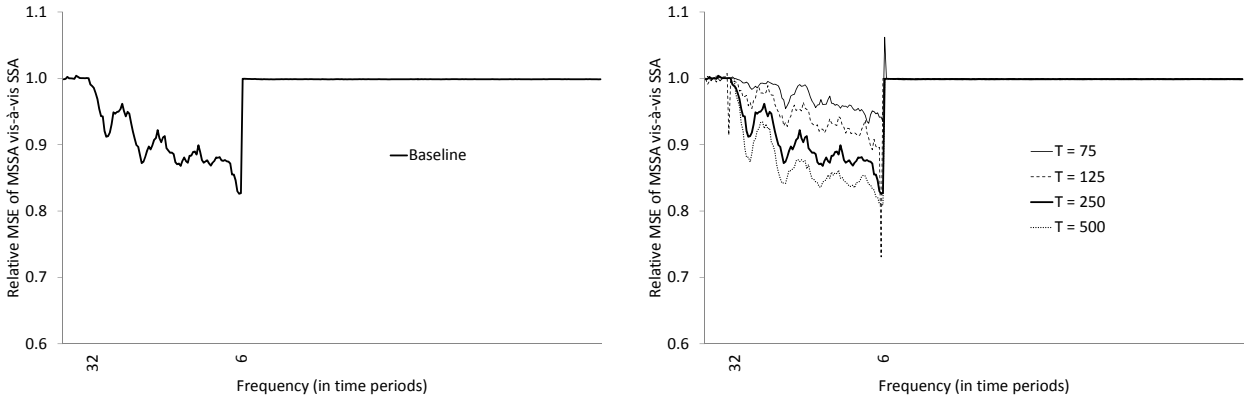
In addition, we also assess the behavior of the SSA based filtering approach when more than one variable is considered. For the sake of parsimony, and because it resembles the case of the empirical application, we consider the two-variable setting. In particular, we consider a second variable, $y_t^{(2)}$, that has the same cyclical component of $y_t^{(1)} := y_t$ but a different noise term. The underlying idea is that we have two variables equally informative about the cycle but perturbed with different noise terms. The multivariate approach intends to retrieve the most of both variables about the cycle, while trying to get rid of the noisy behaviour. Hence, we define a second variable such as

$$y_t^{(2)} = t + \cos(\omega t) + \eta_t,$$

where $0 < \omega < \pi$, $t = 1, \ldots, T$, and $\eta_t$ is Gaussian white noise process with zero mean and variance $\sigma_\eta^2$. To ease the results comparison between the multivariate approach and the univariate counterpart we report the relative MSE, that is, the ratio between the MSE obtained with MSSA and the MSE using the univariate SSA, i.e., we assess the MSE of MSSA vis-à-vis SSA,

$$\text{rMSE}_\omega = \frac{\sum_{t=1}^{T}\sum_{b=1}^{B}\{\widetilde{c}_{t,b} - \cos(\omega t)\}^2}{\sum_{t=1}^{T}\sum_{b=1}^{B}\{\widehat{c}_{t,b} - \cos(\omega t)\}^2}, \quad \omega \in (0, \pi), \tag{17}$$

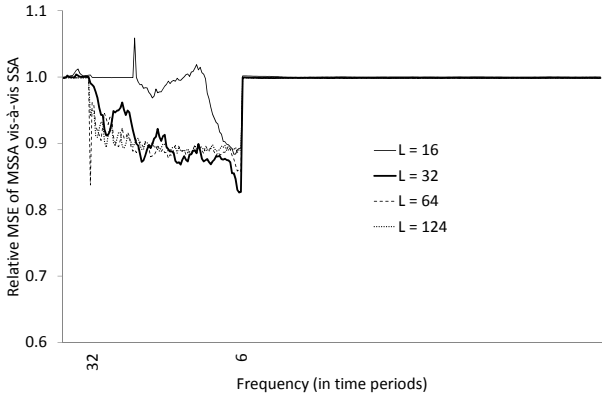
where $\widehat{c}_{t,b}$ and $\widetilde{c}_{t,b}$ respectively denote the SSA and MSSA business cycle indicators. A ratio lower than 1 means that MSSA improves on the univariate approach. From Fig. 3(a) one can see that for the baseline case the relative MSE is lower than 1 for the relevant frequency range with the gains averaging around 10%. In this respect, to obtain further insights on the robustness of the findings we also considered various alternative specifications. By varying the sample size (see Fig. 3(b)), one can conclude that the gains are larger when the sample size increases. Regarding the window length, the results are relatively similar for $L$ equal or larger than 32; see Fig. 3(c). In terms of signal-to-noise ratio, as it can be observed from Fig. 3(d), the gains of using the MSSA vis-à-vis SSA increase when the signal-to-noise ratio decreases. This means that in the presence of noisier series the use of multivariate information becomes relatively more important for signal extraction.



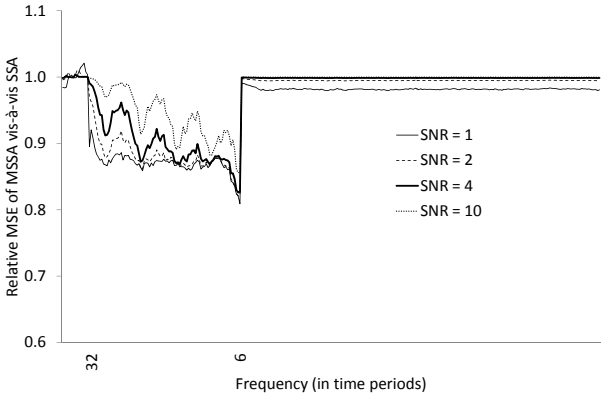

(a)



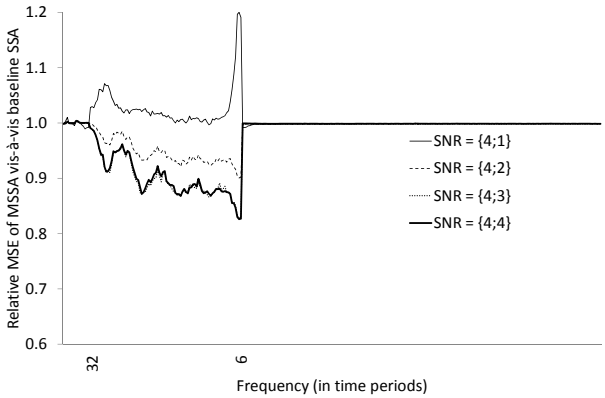

(b)




(c)




(d)




(e)

Figure 3: Simulation results on the multivariate setting. Relative MSE (Mean Square Error) of MSSA vis-à-vis SSA over the spectrum, computed according to Eq. (17). Throughout all figures the bold line is used to represent the baseline specification with $T = 250$, $L = 32$, and with $\sigma_\varepsilon^2$ and $\sigma_\eta^2$ such that the signal-to-noise ratio is 4. Here, $T$ is the sample size, $L$ is the window length, and SNR stands for Signal-to-Noise Ratio.

So far, we have considered that both variables are equally informative about the cyclical component, namely by setting the same signal to noise ratio. In Fig. 3(e), we consider the case when the second variable has a lower signal to noise ratio than the first, which is set to 4 as in the baseline specification. As expected, adding more noise to the second variable reduces the gain of taking it on board when performing the MSSA for extracting the cyclical component.

# 4 Real-time nowcasting the US output gap

## 4.1 Real-time vintages

As the aim is to assess the real-time performance of several alternative methods to extract the cyclical component of GDP, one requires a real-time dataset for the US.

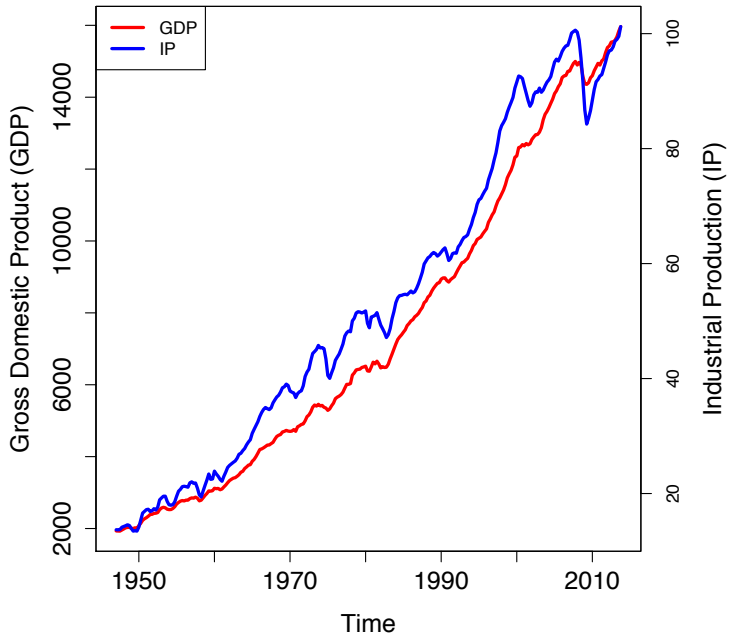



Figure 4: Latest available vintage for the GDP and IP, released on the first quarter of 2014.

In particular, we use the US data set comprising real-time vintages, based on the work of Croushore and Stark (2001), which is maintained by the Federal Reserve Bank of Philadelphia.[6] The sample period

[6] The data are publicly available at:
www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/.

runs from the first quarter of 1947 up to the fourth quarter of 2013. We consider the real-time vintages since the first quarter of 2000, as this is the earliest date for which GDP is available in all subsequent vintages for the whole sample period. In the case of GDP, we use the vintage released at each quarter, which means that for the last quarter it is the first estimate, for the previous quarter it is the second estimate, and so on. In the case of IP we consider the available vintage at the time GDP is released. The period under consideration for real-time evaluation is close to the one in de Carvalho, Rodrigues, and Rua (2012), but extended up to the end of 2013, corresponding to 20% of the sample size; this period encompasses the Great Recession which is by all standards challenging in many dimensions.

## 4.2 Final output gap estimates

In this section, we compute the so-called final output gap estimates, which are based on the latest available vintage (Orphanides & van Norden, 2002); in Fig. 4 we plot the latest available vintage, which for our case corresponds to the one released on the first quarter of 2014. These vintages will then be used to obtain the target output gap for assessing the real-time nowcasting ability of the alternative methods in the next section. Regarding the Hodrick–Prescott filter and the Christiano–Fitzgerald filter we consider the usual parameter values to extract the GDP cyclical component; see Section 3. In the case of the SSA, since we are interested in dynamics of up to 8 years, we set a window length of 32 quarters as in de Carvalho, Rodrigues, and Rua (2012). Regarding the selection of the components in the grouping stage of SSA, we resort to the Fisher $g$-statistic discussed in Section 2.5. Given all the potential components to be considered in the construction of the output gap, we select the components for which the dominant periodicity lies within the standard business cycle frequency range (that is, between 6 and 32 quarters), and which are statistically significant, at the usual 5% significance level, according to the Fisher $g$ test.[7] Once the components are selected, they are aggregated to obtain an output gap measure, by following the steps discussed in Section 2.

[7] We have conducted a sensitivity analysis, considering 1% or 10% significance levels, but the selection of components remains unchanged in what follows.

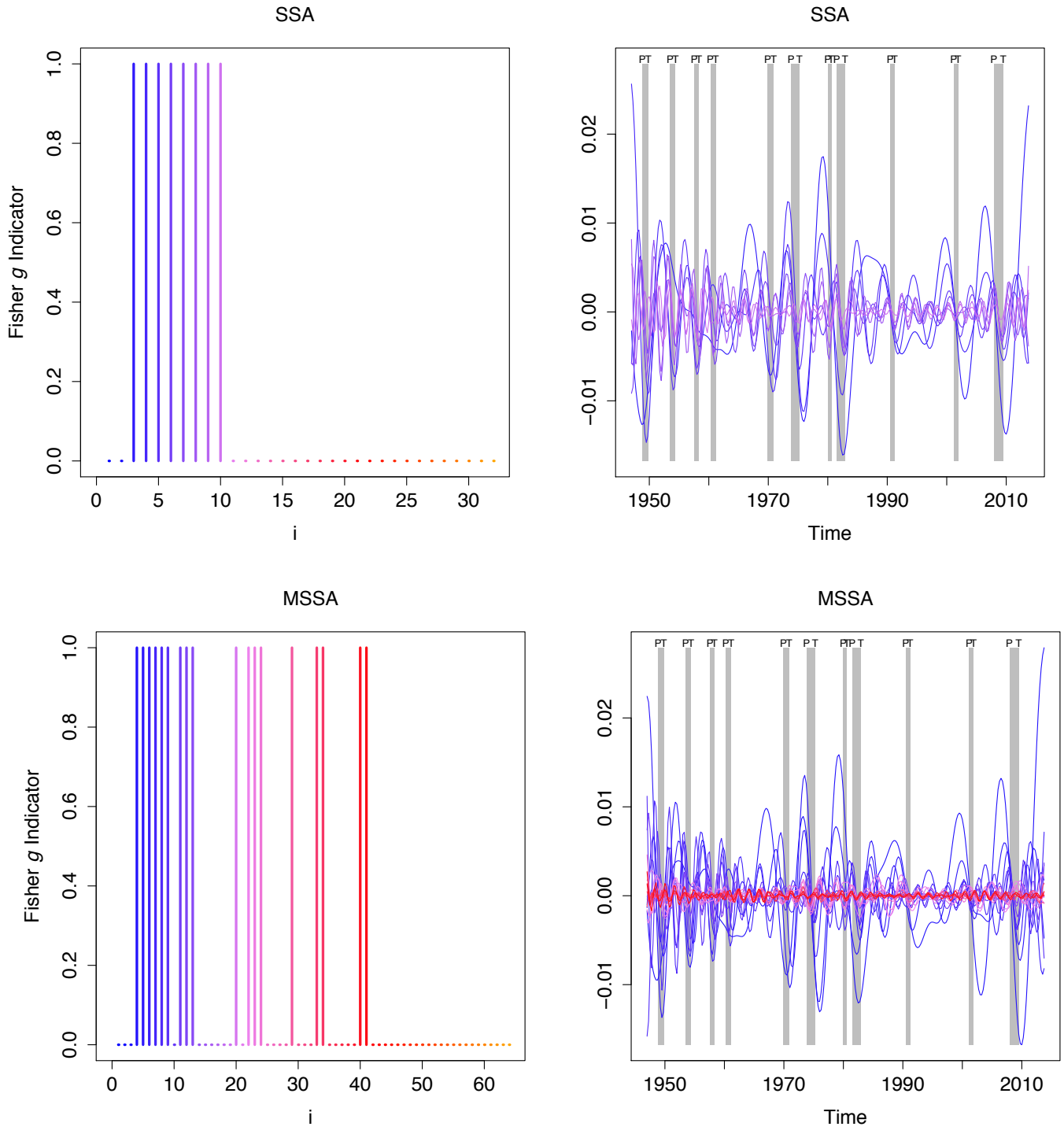



Figure 5: Above: SSA analysis. Below: MSSA analysis. For each of the analysis: on the left, we present the respective comb-plot, as defined in Eq. (15), with the point masses identifying the indices of the components selected according to the approach discussed in Section 2.5; on the right, we plot the principal components used to construct our business cycle indicators, colored according to the same palette as the one used in the comb-plot on the left.

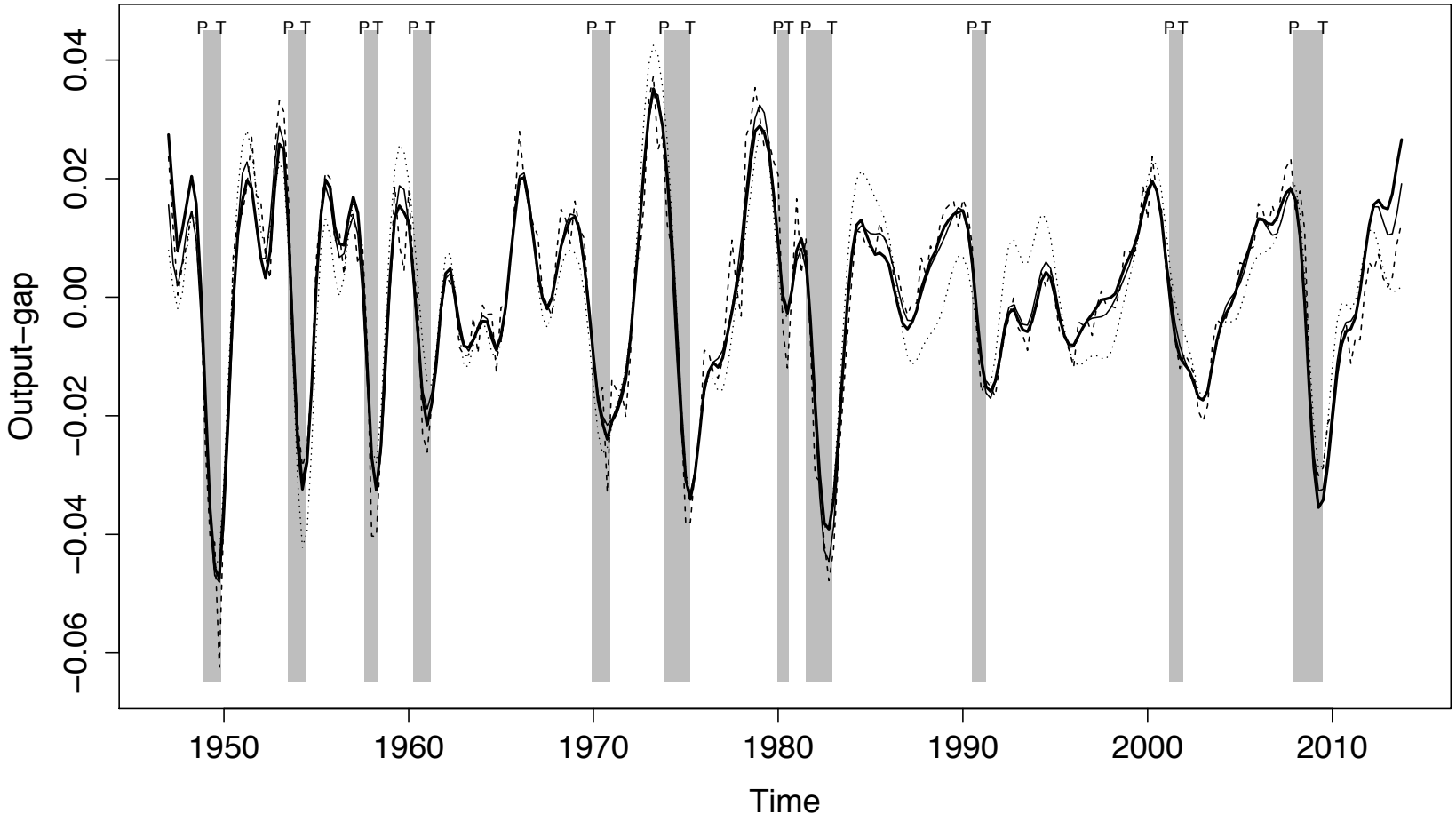



Figure 6: Comparison of output gap estimators: SSA (━), MSSA (—), Hodrick–Prescott (– –), and Christiano–Fitzgerald (· · ·).

The resulting output gap estimates are presented in Fig. 6. All the measures seem to be in accordance with the NBER business cycle chronology and deliver similar qualitative reading concerning the cyclical position of the economy. Note however that, as expected, near the end of the sample there is a higher dispersion of the estimates reflecting the end-of-sample uncertainty. The output gap from the Hodrick–Prescott filter is slightly noisier than the remainder reflecting the fact that it acts as a high-pass filter.

In contrast, the Christiano–Fitzgerald band-pass filter yields a much smoother measure of output gap. In this respect, both SSA and MSSA output gap estimates are also smooth over time, reflecting the criterion adopted in the grouping stage which allows us to discard the trending components and components associated with higher frequencies.

Regarding SSA, the Fisher $g$-statistic led to the selection of the 3rd up to the 10th components, for the construction of the GDP cyclical component, i.e., $S_g = \{3, 4, 5, 6, 7, 8, 9, 10\}$. These almost correspond to the components chosen by de Carvalho, Rodrigues, and Rua (2012), through an heuristic approach which led them to obtain $S = \{3, 4, 5, 6, 7, 8, 9\}$. In practice, the two output gap estimates are nearly indistinguishable graphically; this stems from the fact that the 10th component accounts for a negligible part ($\approx 1\%$) of the variance of the output gap.

In the case of MSSA, from the potential 64 components, 18 have been selected drawing on the Fisher $g$-based approach discussed in Section 2.5; in this case

$$S_g = \{4, 5, 6, 7, 8, 9, 11, 12, 13, 20, 22, 23, 24, 29, 33, 34, 40, 41\},$$

and we summarize this information in the comb-plot in Fig. 5, whose formal definition can be found in Eq. (15).

As can be observed in Fig. 5, in contrast to the case of SSA, the selected components are not sequential in terms of the ordering based on the eingenvalues. In fact, the ordering based on the eigenvalues does not necessarily lead to the most relevant components for the problem at hand. This feature highlights the usefulness of the suggested Fisher $g$-based criterion to identify the components of interest. Naturally, increasing the number of variables makes the practical contribution of using this criterion even more striking.

All in all, the resulting output gap estimates are relatively similar across alternative methods.

Hence, in the next section, we evaluate the information content of the real-time nowcasts for assessing the output gap.

### 4.3 Real-time nowcasting

In this section, we compute the real-time output gap nowcasts based on a recursive estimation exercise, with an expanding sample window, using the real-time vintages of data. In the cases of SSA and MSSA, this also entails the computation of the Fisher $g$ test at each moment in time and corresponding components selection. This truly mimics a real-time scenario in all dimensions. The resulting real-time estimates along with the final output gap estimate for each approach are displayed in Fig. 7.

To evaluate quantitatively the real-time ability of the different methods to nowcast output gap we consider a wide range of performance statistics (see, for example, Orphanides & van Norden, 2002, Marcellino & Musso, 2011). The results are presented in Table 1. In the first column, we report the Mean Absolute Error (MAE), which refers to the average of the absolute difference between the final output gap estimates and the real-time nowcasts. We also present, in the second column, the Root Mean Squared Error (RMSE) which penalizes more larger differences. Both, the MAE and RMSE, are reported in percentage terms. The third column presents the correlation (CORR) between the final and real-time estimates. The next two columns report measures of signal-to-noise of the nowcasts for each method. In particular, SN denotes the ratio of the standard deviation of the final estimate to that of the revision, whereas SNR refers to the ratio of the standard deviation of the final estimate to the root mean square of the revision.[8] In the last two columns we report the sign concordance between the real-time nowcasts and the final estimates. The SIGN-LEV denotes the percentage of times in which the sign of the level of the real-time and final estimates coincide, whereas the SIGN-CH refers to the sign of the changes in output gap estimates.

In terms of the size of the revisions, the Hodrick–Prescott filter seems to perform worse than its competitors, whereas the Christiano–Fitzgerald filter, and the MSSA approach are the top ranked. Among these two, the MSSA method delivers better results than the Christiano–Fitzgerald filter, according to the MAE criterion. In terms of the correlation between the real-time and final estimates,

[8] SN (SNR) are the typical proxies for signal-to-noise ratio in the real-time estimates; cf Orphanides and van Norden (2002, p. 574).

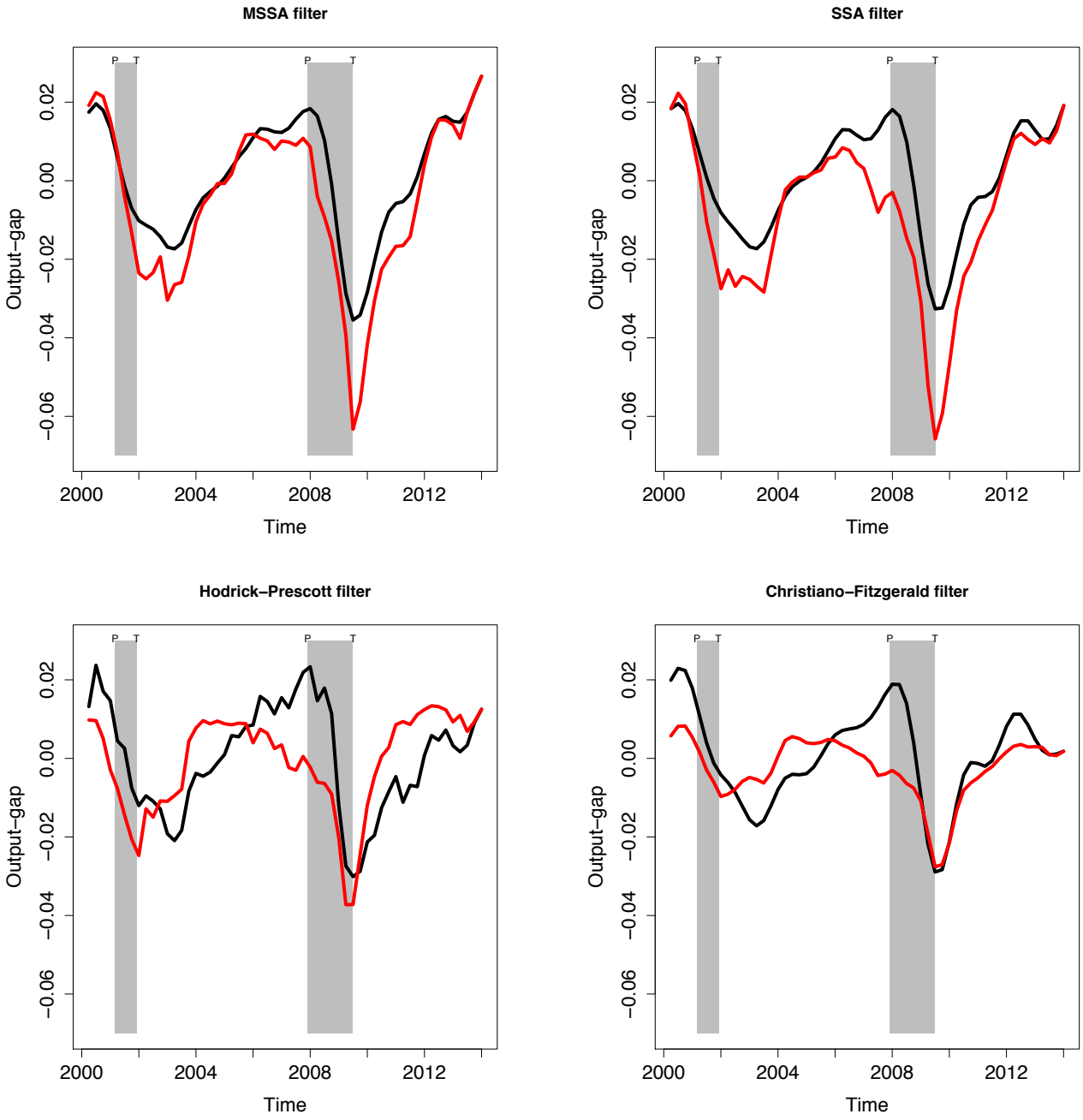



Figure 7: Comparison of real-time estimates (—) and of final estimates (—).

Table 1: Real-time performance evaluation.

| Filter | MAE | RMSE | CORR | SN | SNR | SIGN-LEV | SIGN-CH |
|---|---|---|---|---|---|---|---|
| Hodrick–Prescott | 1.08 | 1.24 | 0.56** | 1.11 | 1.15 | 64.3* | 80.4** |
| Christiano–Fitzgerald | 0.71 | 0.92 | 0.71** | 1.41 | 1.19 | 73.2** | 67.9** |
| SSA | 0.93 | 1.25 | 0.92** | 1.59 | 1.59 | 83.9** | 76.8** |
| MSSA | 0.67** | 0.92** | 0.97** | 2.22 | 2.08 | 92.9** | 80.4** |

NOTE: The top real-time performances according to each measure are identified by the cells in gray. MAE and RMSE are in percentage terms. ** and * denote statistical significance at 1 and 5 % significance levels, respectively. In the cases of MAE and RMSE, we use the modified Diebold and Mariano (1995) test statistic discussed in Harvey, Leybourne, and Newbold (1997) using as benchmark the no output gap estimate. For Pearson correlation coefficient, the usual test statistic is used to assess the null of no correlation. For the sign concordance measures, we consider the test discussed in Hassani, Soofi, and Zhigljavsky (2013).

the Hodrick–Prescott filter ranks last. In contrast, both SSA and MSSA record the highest correlation coefficients, with the latter presenting a correlation close to one. Concerning the signal-to-noise measures, qualitatively similar findings emerge. The Hodrick–Prescott filter records the lowest signal-to-noise whereas there is a striking increase when one considers the MSSA. Although the univariate SSA approach already improves on the Hodrick–Prescott and Christiano–Fitzgerald filters, extending the SSA to the multivariate case results in an even larger increase of the signal-to-noise. Regarding the sign concordance, in terms of the level, the SSA approach outperforms the other filters, with the MSSA standing at the top of the ranking. For the sign concordance in terms of the change, the Christiano–Fitzgerald filter ranks last, whereas the Hodrick–Prescott filter, and MSSA present the best performance. Summing up, for all the performance indicators, the MSSA always ranks first. The SSA approach outperforms, in overall terms, standard filtering techniques, but further gains can still be achieved with the MSSA approach from Section 2.4. Hence, by considering information beyond the one conveyed by GDP, namely the industrial production index, it is possible to improve the real-time performance of the output gap nowcasts in all dimensions. Although it is straightforward to extend the approach in Section 2.4 to a multivariate setting, from an empirical viewpoint it is not clear cut that by enlarging the number of variables considered it will improve the real-time performance. We tried to supplement industrial production with other real-time data, namely non-farm payroll employment

and/or real personal income less transfers, which are also among the series closely monitored by the NBER Business Cycle Dating Committee. However, the results, not reported here, do not reveal any improvement in terms of the performance of the real-time nowcasts.

# 5 Conclusions

This paper explores the performance of SSA-based methods for nowcasting in real-time the US output gap. The assessment in real-time of output gap is of utmost relevance for policymaking, and here we assess the added value of SSA-based nowcasts in a real-life policymaking scenario, by replicating the problem faced by policymakers at the time policy decisions have to be taken. We used real-time vintages, and conducted a recursive study so to evaluate the real-time reliability of our SSA-based approach. For our econometric setting of interest, the preferred specification of our approach consists of a multivariate singular spectrum analysis, where a Fisher $g$ test is used to screen which components—within the standard business cycle range—should be included in the grouping step. Our findings suggest that singular spectrum analysis provides a reliable evaluation of the cyclical position of the US economy in real-time, with the multivariate approach outperforming considerably the univariate counterpart.

Although SSA has been widely applied on many fields of research, there are only a few applications in economics and finance (see, for instance, Hassani, Heravi, & Zhigljavsky, 2009, Patterson, Hassani, Heravi, & Zhigljavsky, 2011, de Carvalho, Rodrigues, & Rua, 2012, Hassani, Heravi, & Zhigljavsky, 2013, Hassani, Soofi, & Zhigljavsky, 2013). We hope that this paper takes another small step in promoting the application of SSA methods in economics, by stressing the resilience of SSA-based approaches to model macroeconomic data. Applied econometric analysis requires the combination of different methodology, and we hope further applied econometricians may consider taking advantage of SSA-based approaches in a near future.

# Acknowledgements

Part of this manuscript was written while M. de. C. was visiting *Banco de Portugal*. The research was partially funded by the Chilean NSF through the Fondecyt project 11121186.

# Appendix

## NBER's Business Cycle Reference Dates

This appendix includes the NBER's business cycle reference dates used in Section 4. This chronology is included here for completeness; the complete chronology can be found at the NBER web site at: www.nber.org/cycles/cyclesmain.html

Table 2: US Business Cycle Reference Dates from Peak to Trough, along with duration of corresponding contractions.

| **Business Cycle Reference Dates** | | |
|---|---|---|
| Peak | Trough | Duration |
| November 1948 (IV) | October 1949 (IV) | 11 |
| July 1953 (II) | May 1954 (II) | 10 |
| August 1957 (III) | April 1958 (II) | 8 |
| April 1960 (II) | February 1961 (I) | 10 |
| December 1969 (IV) | November 1970 (IV) | 11 |
| November 1973 (IV) | March 1975 (I) | 16 |
| January 1980 (I) | July 1980 (III) | 6 |
| July 1981 (III) | November 1982 (IV) | 16 |
| July 1990 (III) | March 1991 (I) | 8 |
| March 2001 (I) | November 2001 (IV) | 8 |
| December 2007 (IV) | June 2009 (II) | 18 |

Here I–IV are used to denote the quarters corresponding to the reference dates; the duration is in months.

# References


A'Hearn, B., & Woitek, U. (2001). More international evidence on the historical properties of business cycles. *Journal of Monetary Economics*, *47*(2), 321–346.

Baxter, M., & King, R. (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics*, *81*(4), 575–593.

Beveridge, S., & Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of business cycle. *Journal of Monetary Economics*, *7*(2), 151–174.

Breitung, J., & Candelon, B. (2006). Testing for short- and long-run causality: A frequency-domain approach. *Journal of Econometrics*, *132*(2), 363–378.

Brockwell, P., & Davis, R. (1991). *Time Series: Theory and Methods*. 2nd Ed. New York: Springer.

Burns, A. F., & Mitchell, W. C. (1946). *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Cassiano, K. M., Júnior, L. A. T., Souza, R. M., Menezes, M. L., Pessanha, J. F. M., & Souza, R. C. (2013). Hydroelectric energy forecast. *International Journal of Energy and Statistics*, *1*(3), 205–214.

Christiano, L., & Fitzgerald, T. (2003). The band-pass filter. *International Economic Review*, *44*(2), 435–465.

Croushore, D., & Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, *105*(1), 111–130.

Croushore, D., & Stark, T. (2003). A real-time data set for macroeconomists: Does the data vintage matter? *The Review of Economics and Statistics*, *85*(3), 605–617.

de Carvalho, M., Rodrigues, P. C., & Rua, A. (2012). Tracking the US business cycle with a singular spectrum analysis. *Economics Letters*, *114*(1), 32–35.

de Carvalho, M., & Rua, A. (2014). Extremal dependence in international output growth: Tales from the tails. *Oxford Bulletin of Economics and Statistics*, *76*(4), 605–620.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 253–263.

Edge, R. M., & Rudd, J. B. (2012). Real-time properties of the Federal Reserve's output gap. *Finance and Economics Discussion Series 2012*(86), *Board of Governors of the Federal Reserve System*.

Elsner, J. B., & Tsonis, A. A. (1996). *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. New York: Plenum.

Fagiolo, G., Napoletano, M., & Roventini, A. (2008). Are output growth-rate distributions fat-tailed? Some evidence from OECD countries. *Journal of Applied Econometrics*, *23*(5), 639–669.

Fisher, R. A. (1929). Tests of significance in harmonic analysis. *Proceedings of the Royal Society of London, Ser.* A, *125*(796), 54–59.

Garratt, A., Mitchell, J., & Vahey, S. P. (2014). Measuring output gap nowcast uncertainty. *International Journal of Forecasting*, *30*, 268–279.

Golyandina, N., Nekrutkin, V., & Zhigljavsky, A. (2001). *Analysis of Time Series Structure: SSA and Related Techniques*. London: Chapman & Hall/CRC.

Harvey, D. I., Leybourne, S. J., & Newbold, P. (1997). Testing equality of prediction mean squared errors. *International Journal of Forecasting*, *13*(2), 281–291.

Hassani, H. (2007). Singular spectrum analysis: Methodology and comparison. *Journal of Data Science*, *5*(2), 239–257.

Hassani, H., Heravi, S., & Zhigljavsky, A. (2009). Forecasting European industrial production with singular spectrum analysis. *International Journal of Forecasting*, *25*(1), 103–118.

Hassani, H., Heravi, S., Brown, G., & Ayoubkhani, D. (2013). Forecasting before, during, and after recession with singular spectrum analysis. *Journal of Applied Statistics*, *40*(10), 2290–2302.

Hassani, H., Heravi, S., & Zhigljavsky, A. (2013). Forecasting UK industrial production with multivariate singular spectrum analysis. *Journal of Forecasting*, *32*(5), 395–408.

Hassani, H., & Mahmoudvand, R. (2013). Multivariate singular spectrum analysis: A general view and new vector forecasting approach. *International Journal of Energy and Statistics*, *1*(1), 55–83.

Hassani, H., Mahmoudvand, R., & Zokaei, M. (2011). Separability and window length in singular spectrum analysis. *Comptes Rendus Mathematique*, *349*(17–18), 987–990.

Hassani, H., Mahmoudvand, R., Zokaei, M., & Ghodsi, M. (2012). On the separability between signal and noise in singular spectrum analysis. *Fluctuation and Noise Letters*, *11*(2), 1250014–1250025.

Hassani, H., Soofi, A. S., & Zhigljavsky, A. (2013). Predicting inflation dynamics with singular spectrum analysis. *Journal of the Royal Statistical Society, Ser.* A, *176*(3), 743–760.

Hassani, H., Webster, A., Silva, E. S., & Heravi, S. (2015). Forecasting U.S. tourist arrivals using optimal singular spectrum analysis. *Tourism Management*, *46*, 322–335.

Hodrick, R. J., & Prescott, E. C. (1997). Postwar US business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, *29*(1), 1–16.

King, R. G., & Rebelo, S. (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, *17*(1), 207–231.

Lemmens, A., Croux, C., & Dekimpe, M. G. (2008). Measuring and testing Granger causality over the spectrum: An application to European production expectation surveys. *International Journal of Forecasting*, *24*(3), 414–431.

Mahmoudvand, R., Najari, N., & Zokaei, M. (2013). On the optimal parameters for reconstruction and forecasting in singular spectrum analysis. *Communication in Statistics: Simulation and Computation*, *42*(4), 860–870.

Marcellino, M., & Musso, A. (2011). The reliability of real-time estimates of the Euro area output gap. *Economic Modelling*, *28*(4), 1842–1856

Mitchell, W. A. (1927). *Business Cycles: The Problem and its Setting.* New York: National Bureau of Economics Research.

Morley, J., & Piger, J. (2012). The asymmetric business cycle. *The Review of Economics and Statistics*, *94*(1), 208–221.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review*, *92*(2), 115–120.

Orphanides, A. (2003a). Monetary policy evaluation with noisy information. *Journal of Monetary Economics*, *50*(3), 605–631.

Orphanides, A. (2003b). The quest for prosperity without inflation. *Journal of Monetary Economics*, *50*(3), 633–663.

Orphanides, A., & van Norden, S. (2002). The unreliability of output gap estimates in real time. *The Review of Economics and Statistics*, *84*(4), 569–583.

Orphanides, A., & Williams, J. C. (2007). Robust monetary policy with imperfect knowledge. *Journal of Monetary Economics*, *54*(5), 1406–1435.

Patterson, K., Hassani, H., Heravi, S., & Zhigljavsky, A. (2011). Multivariate singular spectrum analysis for forecasting revisions to real-time data. *Journal of Applied Statistics*, *38*(10), 2183–2211.

Prescott, E. C. (1986). Theory ahead of business cycle measurement. *Carnegie-Rochester Conference Series on Public Policy*, *25*, 11–66.

Perron, P., & Wada, T. (2009). Let's take a break: Trends and cycles in US real GDP. *Journal of Monetary Economics*, *56*(6), 749–765.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.

Rodrigues, P. C., & de Carvalho, M. (2013). Spectral modeling of time series with missing data. *Applied Mathematical Modelling*, *37*(7), 4676–4684.

Rua, A., & Nunes, L. C. (2005). Coincident and leading indicators for the Euro area: A frequency band approach. *International Journal of Forecasting*, *21*(3), 503–523.

Rudebusch, G. D. (2001). Is the Fed too timid? Monetary policy in an uncertain world. *Review of Economics and Statistics*, *83*(2), 203–217.

Sanei, S., & Hassani, H. (2015). *Singular Spectrum Analysis of Biomedical Signals*. Boca Raton, FL: Chapman & Hall/CRC.

Smets, F. (2002). Output gap uncertainty: Does it matter for the Taylor rule? *Empirical Economics*, *27*(1), 113–129.

Stock, J. H., & Watson, M. W. (1998). Business cycle fluctuations in U.S. macroeconomic time series. In: Taylor, J. B. & Woodford, M. (Eds.), *Handbook of Macroeconomics* (pp. 3–64), Amsterdam: Elsevier.

Stock, J. H., & Watson, M. W. (2005). Understanding changes in international business cycle dynamics. *Journal of the European Economic Association*, *3*(5), 968–1006.

Valle e Azevedo, J. (2011). A multivariate band-pass filter for economic time series. *Journal of the Royal Statistical Society, Ser.* C, *60*(1), 1–30.

Valle e Azevedo, J., Koopman S. J., & Rua, A. (2006). Tracking the business cycle of the Euro area: A multivariate model-based bandpass filter. *Journal of Business and Economic Statistics*, *24*(3), 278–290.

Watson, M. W. (2007). How accurate are real-time estimates of output trends and gaps? *Federal Reserve Bank of Richmond Economic Quarterly*, *93*(2), 143–161.

Yarmohammadi, M. (2011). A filter based Fisher g-test approach for periodicity detection in time series analysis. *Scientific Research and Essays*, *6*(17), 3717–3723.